



---

## Evaluating Data Governance Platforms: A Comprehensive Guide to Selecting the Best Platforms for Your Organization

**Puneet Matai**

Associate Director – Enterprise Data Governance Group  
Santander Bank N.A., USA  
Email: [puneet.matai@gmail.com](mailto:puneet.matai@gmail.com)

---

**Abstract:** This whitepaper guides organizations in selecting the optimal data governance platform by evaluating essential criteria, key features, and performance metrics. It aims to enhance data management, regulatory compliance, and data stewardship. The paper covers governance capabilities, compliance, integration, and scalability. It concludes that choosing a platform with data integration, compliance, and scalability is crucial for effective data governance, recommending careful evaluation and consideration of these factors.

**Keywords:** Data Governance, Enterprise Data Strategy, Compliance, Data Integration, Data Security, GDPR, CCPA, Privacy, Data Management, Data Catalog, Metadata, Data Lineage

---

### Introduction

The emergence of ecosystems centred around platform-based infrastructures highlights the transformative potential of data in creating value. Organizations are faced with the challenge of balancing these opposing interests. Effective governance structures and models ensure that data sharing occurs in a controlled and secure manner [1].

On this note, choosing the right data governance platform becomes crucial for effective data management and regulatory compliance. A well-selected platform will ensure accurate data stewardship, enhance visibility into data flows, and comply with regulations like GDPR and CCPA.

The whitepaper aims to guide organizations in selecting the most suitable data governance platform by providing the evaluation of leading platforms. The evaluation covers essential criteria, key features, performance metrics, and specific platforms to facilitate an informed decision in data governance selection.

### Essential Criteria for Evaluation

#### Governance Capabilities

The recent surge in data catalogs is driven by the increasing use of analytics across all departments in large companies, regardless of data literacy. A data catalog serves several key governance capabilities:

**Data Stewardship:** Data stewards are responsible for the implementation and enforcement of data governance policies and standards within their organization by assigning roles and responsibilities. NOAA's OneStop Metadata team uses MM-Stew assessments to rate Earth Science data products integrating these ratings into ISO 19115 metadata for improved discoverability [2].

**Data Lineage:** It tracks data's lifecycle, transformations, and movement to ensure clarity, and efficient management, and prevent redundant workloads. There are automated data lineage tools which map relationships and conduct impact analysis to enhance root cause analysis and data visibility [3].

**Metadata Management:** It stores detailed information about data assets, including usage and relationships. It reduces the likelihood of compliance breaches due to incorrect classification or inappropriate sharing.



**Data Quality Management:** Implementing quality practices helps in identifying and correcting data errors promptly. This reduces the risk of data-related issues.

#### **Compliance and Security Features**

When evaluating data governance platforms, it is essential to consider compliance and security features such as: **Adherence to Regulations such as GDPR, CCPA:** General Data Protection Regulation (GDPR) and California Consumer Privacy Act (CCPA) impose rules and regulations on how organizations handle personal data. A platform that ensures compliance helps mitigate these risks.

**Data Protection:** These regulations grant individual rights over their data, such as the right to access, rectify, and delete their information. A platform must facilitate these to ensure respect for individual rights.

**Access Controls:** It often includes logging and monitoring capabilities that provide detailed audit trails of data access and usage. Granular access controls minimize risks of unauthorised access.

#### **Integration and Compatibility**

A data governance platform must integrate seamlessly with an organization's existing systems and data sources. Integration capabilities enable the consolidation of data from various sources into a single platform.

#### **Scalability and Performance**

As organizations grow, their data volumes increase. A scalable platform can accommodate this growth without performance degradation. Performance benchmarks are important to ensure that the platform can conduct quick data processing, low latency, and high throughput.

### **Key Features to Consider**

#### **User Interface and Usability**

Ease of use, customization options, and user training are critical for ensuring that the platform is accessible and adaptable to diverse user needs. A user-friendly interface encourages widespread adoption, while customization enhances flexibility. Comprehensive training programs ensure users can effectively leverage the platform's capabilities.

#### **Automation and Workflow Management**

Automation and workflow management streamline data governance tasks, reducing manual effort and enhancing efficiency. These features enable consistent application of governance policies and facilitate the management of complex workflows.

#### **Reporting and Analytics**

Reporting and analytics capabilities are essential for generating actionable insights, creating dashboards, and maintaining audit trails. These features enable organizations to monitor compliance, track data quality, and support decision-making.

#### **Collaboration and Support**

Effective collaboration tools and quality vendor support are vital for team coordination and problem resolution. These features ensure that users can work together, share knowledge, and receive timely assistance.

### **Performance Metrics**

#### **Efficiency Metrics**

The speed at which a platform processes data is crucial for maintaining operational efficiency and meeting business needs. Fast data processing allows for real-time analytics, timely decision-making, and improved productivity. Assess the platform's ability to handle large volumes of data quickly without performance degradation

#### **Accuracy and Reliability**

Users need to access data and receive responses promptly to perform their tasks effectively. Measure the time it takes for the platform to respond to user queries and commands. Further, shorter response times indicate a more efficient and user-friendly system.

#### **Cost vs Value**

Total Cost of Ownership (TCO): Understanding the Total Cost of Ownership, which includes initial costs, ongoing maintenance, and operational expenses are crucial for budgeting and financial planning. The research



by [4] indicates that TCO is considered a tool which supports the development of Product-Service Systems (PSS).

**Return on Investment (ROI):** ROI measures the financial return from the investment in the data governance platform. A higher ROI indicates that the platform delivers substantial benefits relative to its cost.

**Value for Money:** Beyond just the cost, it is important to consider the value the platform provides. This will include the features, capabilities, support, and long-term benefits relative to its price.

### Market Leading Platforms

This section provides an overview of the leading data governance platforms and compares them based on essential criteria:

#### Collibra

**Description:** Collibra is a data intelligence platform which delivers flexible governance, continuous quality, and built-in privacy to help organizations effectively manage and utilize trusted data.

**Features:** Artificial Intelligence (AI) Governance, Data Reliability, Data Cloud Migration, Data Governance Strategy, Data Reliability, and Data Sharing [5].

#### Alation

**Description:** Alation empowers enterprise innovation by providing a data catalog which promotes data-driven collaboration [6]. It enables organizations to create thriving data cultures and implement data for insights.

**Features:** Active Data Governance, AI Governance, Data Modernization, Data Quality, Metadata Management, Privacy-Risk-Compliance [6]

#### Informatica

**Description:** Informatica provides a comprehensive suite of data management solutions, including data governance, data quality, master data management, and data integration.

**Features:** Data Catalog, Cloud Modernization, SAP Modernization, MDM Modernization, API & App Integration, governance etc [7].

#### IBM Watson Knowledge Catalog

**Description:** It is a cloud-based data cataloging and governance solution that implements AI to automate data discovery, classification, and governance [8]. It integrates with various IBM data solutions to provide a cohesive data governance framework.

**Features:** Data Discovery, Data Catalog, Data Lineage, Metadata Management, Policy Management.

**Here's a brief comparison of the platforms:**

| Criteria                     | Collibra  | Alation  | Informatica   | IBM Watson Knowledge Catalog                  |
|------------------------------|---|--|---|---|
| Governance Capabilities      | Comprehensive stewardship and quality management.   | Advanced cataloging and stewardship                | Extensive governance and quality management.              | AI-driven discovery and classification.       |
| Compliance and Security      | GDPR/CCPA Compliance                                | Data access controls                               | Integrated compliance and data protection                 | AI-driven policy management.                  |
| Accuracy and Reliability     | Low error rates, High data accuracy and consistency | Reliable data accuracy, Effective error management | High accuracy, strong data consistency and error handling | High accuracy, AI-driven validation.          |
| Scalability and Performance  | High scalability for large data volumes             | Scalable architecture                              | Highly scalable for enterprise-level governance           | Scalable cloud-based performance              |
| User Interface and Usability | Intuitive, customizable interface                   | User-friendly with strong collaboration features   | Comprehensive UI with extensive customization             | Modern, intuitive interface                   |
| Efficiency Metrics           | High-speed processing, Quick response times         | Efficient processing and fast response times.      | Exceptional processing speeds, High responsiveness        | Fast data processing, AI-optimized efficiency |



## Evaluation Dimensions

When evaluating data governance platforms, it's crucial to assess dimensions beyond core functionalities. Key dimensions include:

### Vendor Reputation and Track Record

The platform's standing in the market can indicate its reliability and the level of trust it has earned from users. For example, Collibra and Informatica are recognized leaders in the data governance space, with a significant share of the global market.

### Flexibility and Customization

Flexibility and customization capabilities are vital for tailoring the platform to specific organizational requirements, including workflow adjustments, data governance policies, and integration with existing systems. Platforms should support technical customizations such as API integrations, custom reporting, and workflow automation to align with specific business processes.

For example, Informatica supports extensive API integrations, with over 100 connectors available for integration with various data sources and applications.

### Future-Proofing

Evaluating a platform's commitment to innovation and its development roadmap helps assess its ability to incorporate new technologies and address emerging trends. A platform's ability to adapt to emerging trends, such as big data, cloud computing, and AI, ensures it remains relevant as technology evolves.

For example, IBM Watson Knowledge implements AI for data management and discovery, positioning it for future advancements.

## Pitfalls And Challenges

When evaluating and implementing data governance platforms, several common pitfalls and challenges can hinder successful adoption and utilization. A survey by [9] has shown that across a broad spectrum of companies, around 43% experienced a data breach during 2013 and it increased by 10% from prior years.

The common pitfalls and challenges are:

### Overlooking Integration Challenges

Failure to thoroughly assess integration capabilities can lead to significant issues during implementation, such as data silos, workflow disruptions, and compatibility problems. Organizations may face delays, increased costs, and operational inefficiencies if the platform does not seamlessly integrate with the current infrastructure.

Consider the evidence for data definitions, protection, quality assurance, and sharing protocols.

### Underestimating User Training Needs

Many organizations overlook the importance of comprehensive user training and support. Insufficient training can lead to suboptimal use of the platform, errors, and resistance to adoption.

Evaluate data security measures, formal policies, and staff responsibilities.

## Conclusion And Recommendations

Selecting the right data governance is important for effective data management, compliance, and operational efficiency.

By evaluating governance capabilities, compliance features, integration, scalability, and performance metrics, organizations can make informed decisions that align with their specific needs and future growth.

### It is recommended that:

**Prioritize Integration and Compliance:** Choose a platform that integrates with existing systems seamlessly and complies with GDPR and CCPA.

**Evaluate Scalability and Performance:** Select a platform that offers high scalability and performance metrics to handle growing data volumes efficiently.

## References

- [1]. D. Lis and B. Otto, "Data Governance in Data Ecosystems - Insights from Organizations.," Americas Conference on Information Systems, Jan. 2020. (accessed Oct. 15, 2020).



- [2]. G. Peng, "The State of Assessing Data Stewardship Maturity – An Overview," *Data Science Journal*, vol. 17, 2018, doi: <https://doi.org/10.5334/dsj-2018-007>. (accessed Oct 15, 2020).
- [3]. O. Hýbl, "Data Lineage Analysis of Frameworks with Complex Interaction Patterns," Cuni.cz, 2020, doi: <http://hdl.handle.net/20.500.11956/147591>. (accessed Oct. 15, 2020).
- [4]. I. Roda, M. Macchi, and S. Albanese, "Building a Total Cost of Ownership model to support manufacturing asset lifecycle management," *Production Planning & Control*, vol. 31, no. 1, pp. 19–37, Jun. 2019, doi: <https://doi.org/10.1080/09537287.2019.1625079>. (accessed Oct. 15, 2020).
- [5]. Collibra, "Discover the Data Intelligence Company | Collibra," Collibra.com, 2020. <https://www.collibra.com/us/en/company/who-we-are?> (accessed Oct. 15, 2020).
- [6]. "Enterprise Data Catalog & Data Governance," Alation. <https://www.alation.com/> (accessed Oct. 15, 2020).
- [7]. "Informatica," Informatica. <https://www.informatica.com/in/> (accessed Oct. 18, 2020).
- [8]. IBM, "IBM Knowledge Catalog," [www.ibm.com](http://www.ibm.com). <https://www.ibm.com/products/knowledge-catalog> (accessed Oct. 18, 2020).
- [9]. D. Mauzy, B. Bull, and T. Gould, "Avoid the Pitfalls: Benefits of Formal Part C Data System Governance Data Governance Lessons Learned from Other Industries: Keep Data Secure." [Online]. Available: <https://dasycenter.org/wp-content/uploads/2016/05/DGMY3-1-WhyDGMMatters.pdf> (accessed Oct. 18, 2020).

